

#4

OIPE

0590
06107

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/782,051

DATE: 06/20/2001

TIME: 14:17:13

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

ENTERED

3 <110> APPLICANT: Hart, Derek N J
 5 <120> TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
 6 (AHCY) type activity
 8 <130> FILE REFERENCE: 24305 MRB
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/782,051
 C--> 11 <141> CURRENT FILING DATE: 2001-02-14
 13 <150> PRIOR APPLICATION NUMBER: PCT/NZ97/00133
 14 <151> PRIOR FILING DATE: 1997-10-06
 16 <150> PRIOR APPLICATION NUMBER: NZ 299507
 17 <151> PRIOR FILING DATE: 1996-10-04
 19 <160> NUMBER OF SEQ ID NOS: 2
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 2563
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Homo sapiens
 28 <220> FEATURE:
 29 <221> NAME/KEY: CDS
 30 <222> LOCATION: (3)..(1847)
 31 <223> OTHER INFORMATION: Open reading frame extends without a stop codon
 32 for the full 5' nucleotide sequence. The
 33 initiation codon has yet to be identified.
 35 <400> SEQUENCE: 1
 36 gg cgc ggg cag gtc gga gct cgg agc tgc tgc ttc tgg ttc tct tgt 47
 37 Arg Gly Gln Val Gly Ala Arg Ser Cys Cys Phe Trp Phe Ser Cys
 38 1 5 10 15
 40 ggc cac cgt cgc tgt ccg gct gcc ttg ggc tgc cga aca gac aag gcg 95
 41 Gly His Arg Arg Cys Pro Ala Ala Leu Gly Cys Arg Thr Asp Lys Ala
 42 20 25 30
 44 tgg gcc aca gca cct cag aag ccg acg cag ctc gac gca ggg gcc ggc 143
 45 Trp Ala Thr Ala Pro Gln Lys Pro Thr Gln Leu Asp Ala Gly Ala Gly
 46 35 40 45
 48 agg agg gtg ggc gat cgc gtg tcg gag ggc gcc gcg cgg gca ggc ggc 191
 49 Arg Arg Val Gly Asp Arg Val Ser Glu Gly Ala Ala Arg Ala Gly Gly
 50 50 55 60
 52 cgg gcg cca gag ggg gaa aga ggc ggg ggc ggc ggg tca gcc gct ggc 239
 53 Arg Ala Pro Glu Gly Glu Arg Gly Gly Gly Gly Gly Ser Ala Ala Gly
 54 65 70 75
 56 cgg gcc ggc ggg gga atg tcg atg cct gac gcg atg ccg ctg ccc ggc 287
 57 Arg Ala Gly Gly Gly Met Ser Met Pro Asp Ala Met Pro Leu Pro Gly
 58 80 85 90 95
 60 gtc ggg gag gag ctg aag cag gcc aag gag atc gag gac gcc gag aag 335
 61 Val Gly Glu Glu Leu Lys Gln Ala Lys Glu Ile Glu Asp Ala Glu Lys
 62 100 105 110
 64 tac tcc ttc atg gcc acc gtc acc aag gcg ccc aag aag caa atc cag 383
 65 Tyr Ser Phe Met Ala Thr Val Thr Lys Ala Pro Lys Lys Gln Ile Gln
 66 115 120 125

RAW SEQUENCE LISTING

DATE: 06/20/2001

PATENT APPLICATION: US/09/782,051

TIME: 14:17:13

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

68	ttt gct gat gac atg cag gag ttc acc aaa ttc ccc acc aaa act ggc	431
69	Phe Ala Asp Asp Met Gln Glu Phe Thr Lys Phe Pro Thr Lys Thr Gly	
70	130 135 140	
72	cga aga tct ttg tct cgc tcg atc tca cag tcc tcc act gac agc tac	479
73	Arg Arg Ser Leu Ser Arg Ser Ile Ser Gln Ser Ser Thr Asp Ser Tyr	
74	145 150 155	
76	agt tca gct gca tcc tac aca gat agc tct gat gat gag gtt tct ccc	527
77	Ser Ser Ala Ala Ser Tyr Thr Asp Ser Ser Asp Asp Glu Val Ser Pro	
78	160 165 170 175	
80	cga gag aag cag caa acc aac tcc aag ggc agc agc aat ttc tgt gtg	575
81	Arg Glu Lys Gln Gln Thr Asn Ser Lys Gly Ser Ser Asn Phe Cys Val	
82	180 185 190	
84	aag aac atc aag cag gca gaa ttt gga cgc cgg gag att gag att gca	623
85	Lys Asn Ile Lys Gln Ala Glu Phe Gly Arg Arg Glu Ile Glu Ile Ala	
86	195 200 205	
88	gag caa gac atg tct gct ctg att tca ctc agg aaa cgt gct cag ggg	671
89	Glu Gln Asp Met Ser Ala Leu Ile Ser Leu Arg Lys Arg Ala Gln Gly	
90	210 215 220	
92	gag aag ccc ttg gct ggt gct aaa ata gtg ggc tgt aca cac atc aca	719
93	Glu Lys Pro Leu Ala Gly Ala Lys Ile Val Gly Cys Thr His Ile Thr	
94	225 230 235	
96	gcc cag aca gcg gtg ttg att gag aca ctc tgt gcc ctg ggg gct cag	767
97	Ala Gln Thr Ala Val Leu Ile Glu Thr Leu Cys Ala Leu Gly Ala Gln	
98	240 245 250 255	
100	tgc cgc tgg tct gct tgt aac atc tac tca act cag aat gaa gta gct	815
101	Cys Arg Trp Ser Ala Cys Asn Ile Tyr Ser Thr Gln Asn Glu Val Ala	
102	260 265 270	
104	gca gca ctg gct gag gct gga gtt gca gtg ttc gct tgg aag ggc gag	863
105	Ala Ala Leu Ala Glu Ala Gly Val Ala Val Phe Ala Trp Lys Gly Glu	
106	275 280 285	
108	tca gaa gat gac ttc tgg tgg tgt att gac cgc tgt gtg aac atg gat	911
109	Ser Glu Asp Asp Phe Trp Trp Cys Ile Asp Arg Cys Val Asn Met Asp	
110	290 295 300	
112	ggg tgg cag gcc aac atg atc ctg gat gat ggg gga gac tta acc cac	959
113	Gly Trp Gln Ala Asn Met Ile Leu Asp Asp Gly Gly Asp Leu Thr His	
114	305 310 315	
116	tgg gtt tat aag aag tat cca aac gtg ttt aag aag atc cga ggc att	1007
117	Trp Val Tyr Lys Lys Tyr Pro Asn Val Phe Lys Lys Ile Arg Gly Ile	
118	320 325 330 335	
120	gtg gaa gag agc gtg act ggt gtt cac agg ctg tat cag ctc tcc aaa	1055
121	Val Glu Glu Ser Val Thr Gly Val His Arg Leu Tyr Gln Leu Ser Lys	
122	340 345 350	
124	gct ggg aag ctc tgt gtt ccg gcc atg aac gtc aat gat tct gtt acc	1103
125	Ala Gly Lys Leu Cys Val Pro Ala Met Asn Val Asn Asp Ser Val Thr	
126	355 360 365	
128	aaa cag aag ttt gat aac ttg tac tgc tgc cga gaa tcc att ttg gat	1151
129	Lys Gln Lys Phe Asp Asn Leu Tyr Cys Cys Arg Glu Ser Ile Leu Asp	
130	370 375 380	
132	ggc ctg aag agg acc aca gat gtg atg ttt ggt ggg aaa caa gtg gtg	1199

RAW SEQUENCE LISTING

DATE: 06/20/2001

PATENT APPLICATION: US/09/782,051

TIME: 14:17:13

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

```

133 Gly Leu Lys Arg Thr Thr Asp Val Met Phe Gly Gly Lys Gln Val Val
134      385      390      395
136 gtg tgt ggc tat ggt gag gta ggc aag ggc tgc tgt gct gct ctc aaa 1247
137 Val Cys Gly Tyr Gly Glu Val Gly Lys Gly Cys Cys Ala Ala Leu Lys
138 400      405      410      415
140 gct ctt gga gca att gtc tac att acc gaa atc gac ccc atc tgt gct 1295
141 Ala Leu Gly Ala Ile Val Tyr Ile Thr Glu Ile Asp Pro Ile Cys Ala
142      420      425      430
144 ctg cag gcc tgc atg gat ggg ttc agg gtg gta aag cta aat gaa gtc 1343
145 Leu Gln Ala Cys Met Asp Gly Phe Arg Val Val Lys Leu Asn Glu Val
146      435      440      445
148 atc cgg caa gtc gat gtc gta ata act tgc aca gga aat aag aat gta 1391
149 Ile Arg Gln Val Asp Val Val Ile Thr Cys Thr Gly Asn Lys Asn Val
150      450      455      460
152 gtg aca cgg gag cac ttg gat cgc atg aaa aac agt tgt atc gta tgc 1439
153 Val Thr Arg Glu His Leu Asp Arg Met Lys Asn Ser Cys Ile Val Cys
154      465      470      475
156 aat atg ggc cac tcc aac aca gaa atc gat gtg acc agc ctc cgc act 1487
157 Asn Met Gly His Ser Asn Thr Glu Ile Asp Val Thr Ser Leu Arg Thr
158 480      485      490      495
160 ccg gag ctg acg tgg gag cga gta cgt tct cag gtg gac cat gtc atc 1535
161 Pro Glu Leu Thr Trp Glu Arg Val Arg Ser Gln Val Asp His Val Ile
162      500      505      510
164 tgg cca gat ggc aaa cga gtt gtc ctc ctg gca gag ggt cgt cta ctc 1583
165 Trp Pro Asp Gly Lys Arg Val Val Leu Leu Ala Glu Gly Arg Leu Leu
166      515      520      525
168 aat ttg agc tgc tcc aca gtt ccc acc ttt gtt ctg tcc atc aca gcc 1631
169 Asn Leu Ser Cys Ser Thr Val Pro Thr Phe Val Leu Ser Ile Thr Ala
170      530      535      540
172 aca aca cag gct ttg gca ctg ata gaa ctc tat aat gca ccc gag ggg 1679
173 Thr Thr Gln Ala Leu Ala Leu Ile Glu Leu Tyr Asn Ala Pro Glu Gly
174      545      550      555
176 cga tac aag cag gat gtg tac ttg ctt cct aag aaa atg gat gaa tac 1727
177 Arg Tyr Lys Gln Asp Val Tyr Leu Leu Pro Lys Lys Met Asp Glu Tyr
178 560      565      570      575
180 gtt gcc agc ttg cat ctg cca tca ttt gat gcc cac ctt aca gag ctg 1775
181 Val Ala Ser Leu His Leu Pro Ser Phe Asp Ala His Leu Thr Glu Leu
182      580      585      590
184 aca gat gac caa gca aaa tat ctg gga ctc aac aaa aat ggg cca ttc 1823
185 Thr Asp Asp Gln Ala Lys Tyr Leu Gly Leu Asn Lys Asn Gly Pro Phe
186      595      600      605
188 aaa cct aat tat tac aga tac taa tggaccatac taccaaggac cagtccacct 1877
189 Lys Pro Asn Tyr Tyr Arg Tyr
W--> 190      610      615
192 gaaccacaca ctctaaagaa atatttttta agataacttt tattttcttc ttactccttt 1937
194 cctcttgatt tttttcctat aatttcattc ttgttttttc atctcattat ccaagttctg 1997
196 cagaccacac aggaacttgc ttcattggctc tttagatgaa atagaagttc agggtccttc 2057
198 actctagtca ctaaagaagg attttactcc cccagccag aaaggtgatt cttctcttta 2117
200 ccatttctgg ggacttttagt cttaattagg taccttatta acaggaaatg ctaaggtacc 2177

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/782,051

DATE: 06/20/2001

TIME: 14:17:13

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

```

202 ttctctgtgg aacaatctgc aatgtctaaa tcgccttaaa agagcccatt tcttagctgc 2237
204 tgaaatcagt gctctttcac ttcttcagag aagcagggat ggtacctacc cggcaggtag 2297
206 gttagatgtg ggtggtgcat gttaatttcc cttagaagtt ccaagccctg tttcctgcgt 2357
208 aaaggtggta tgtccagttc agagatgtgt ataatgagca tggcttggtta agatcaggag 2417
210 gccacttgg atttatagta tagcccttcc tccactccca ccagacttgc tcatttttcg 2477
212 agttttttaac tagactacac tctatttgag tttaattttg tcctctagga tttatttctg 2537
214 ttgtccaaaa aaaaaaaaaa aaaaaa 2563
217 <210> SEQ ID NO: 2
218 <211> LENGTH: 614
219 <212> TYPE: PRT
220 <213> ORGANISM: Homo sapiens
222 <400> SEQUENCE: 2
223 Arg Gly Gln Val Gly Ala Arg Ser Cys Cys Phe Trp Phe Ser Cys Gly
224 1 5 10 15
226 His Arg Arg Cys Pro Ala Ala Leu Gly Cys Arg Thr Asp Lys Ala Trp
227 20 25 30
229 Ala Thr Ala Pro Gln Lys Pro Thr Gln Leu Asp Ala Gly Ala Gly Arg
230 35 40 45
232 Arg Val Gly Asp Arg Val Ser Glu Gly Ala Ala Arg Ala Gly Gly Arg
233 50 55 60
235 Ala Pro Glu Gly Glu Arg Gly Gly Gly Gly Gly Ser Ala Ala Gly Arg
236 65 70 75 80
238 Ala Gly Gly Gly Met Ser Met Pro Asp Ala Met Pro Leu Pro Gly Val
239 85 90 95
241 Gly Glu Glu Leu Lys Gln Ala Lys Glu Ile Glu Asp Ala Glu Lys Tyr
242 100 105 110
244 Ser Phe Met Ala Thr Val Thr Lys Ala Pro Lys Lys Gln Ile Gln Phe
245 115 120 125
247 Ala Asp Asp Met Gln Glu Phe Thr Lys Phe Pro Thr Lys Thr Gly Arg
248 130 135 140
250 Arg Ser Leu Ser Arg Ser Ile Ser Gln Ser Ser Thr Asp Ser Tyr Ser
251 145 150 155 160
253 Ser Ala Ala Ser Tyr Thr Asp Ser Ser Asp Asp Glu Val Ser Pro Arg
254 165 170 175
256 Glu Lys Gln Gln Thr Asn Ser Lys Gly Ser Ser Asn Phe Cys Val Lys
257 180 185 190
259 Asn Ile Lys Gln Ala Glu Phe Gly Arg Arg Glu Ile Glu Ile Ala Glu
260 195 200 205
262 Gln Asp Met Ser Ala Leu Ile Ser Leu Arg Lys Arg Ala Gln Gly Glu
263 210 215 220
265 Lys Pro Leu Ala Gly Ala Lys Ile Val Gly Cys Thr His Ile Thr Ala
266 225 230 235 240
268 Gln Thr Ala Val Leu Ile Glu Thr Leu Cys Ala Leu Gly Ala Gln Cys
269 245 250 255
271 Arg Trp Ser Ala Cys Asn Ile Tyr Ser Thr Gln Asn Glu Val Ala Ala
272 260 265 270
274 Ala Leu Ala Glu Ala Gly Val Ala Val Phe Ala Trp Lys Gly Glu Ser
275 275 280 285
277 Glu Asp Asp Phe Trp Trp Cys Ile Asp Arg Cys Val Asn Met Asp Gly

```

RAW SEQUENCE LISTING

DATE: 06/20/2001

PATENT APPLICATION: US/09/782,051

TIME: 14:17:13

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

278	290	295	300
280 Trp	Gln Ala Asn Met Ile	Leu Asp Asp Gly Gly	Asp Leu Thr His Trp
281 305		310	315
283 Val	Tyr Lys Lys Tyr Pro	Asn Val Phe Lys Lys	Ile Arg Gly Ile Val
284		325	330
286 Glu	Glu Ser Val Thr Gly	Val His Arg Leu Tyr	Gln Leu Ser Lys Ala
287		340	345
289 Gly	Lys Leu Cys Val Pro	Ala Met Asn Val Asn	Asp Ser Val Thr Lys
290		355	360
292 Gln	Lys Phe Asp Asn Leu	Tyr Cys Cys Arg Glu	Ser Ile Leu Asp Gly
293		370	375
295 Leu	Lys Arg Thr Thr Asp	Val Met Phe Gly Gly	Lys Gln Val Val Val
296 385		390	395
298 Cys	Gly Tyr Gly Glu Val	Gly Lys Gly Cys Cys	Ala Ala Leu Lys Ala
299		405	410
301 Leu	Gly Ala Ile Val Tyr	Ile Thr Glu Ile Asp	Pro Ile Cys Ala Leu
302		420	425
304 Gln	Ala Cys Met Asp Gly	Phe Arg Val Val Lys	Leu Asn Glu Val Ile
305		435	440
307 Arg	Gln Val Asp Val Val	Ile Thr Cys Thr Gly	Asn Lys Asn Val Val
308		450	455
310 Thr	Arg Glu His Leu Asp	Arg Met Lys Asn Ser	Cys Ile Val Cys Asn
311 465		470	475
313 Met	Gly His Ser Asn Thr	Glu Ile Asp Val Thr	Ser Leu Arg Thr Pro
314		485	490
316 Glu	Leu Thr Trp Glu Arg	Val Arg Ser Gln Val	Asp His Val Ile Trp
317		500	505
319 Pro	Asp Gly Lys Arg Val	Val Leu Leu Ala Glu	Gly Arg Leu Leu Asn
320		515	520
322 Leu	Ser Cys Ser Thr Val	Pro Thr Phe Val Leu	Ser Ile Thr Ala Thr
323		530	535
325 Thr	Gln Ala Leu Ala Leu	Ile Glu Leu Tyr Asn	Ala Pro Glu Gly Arg
326 545		550	555
328 Tyr	Lys Gln Asp Val Tyr	Leu Leu Pro Lys Lys	Met Asp Glu Tyr Val
329		565	570
331 Ala	Ser Leu His Leu Pro	Ser Phe Asp Ala His	Leu Thr Glu Leu Thr
332		580	585
334 Asp	Asp Gln Ala Lys Tyr	Leu Gly Leu Asn Lys	Asn Gly Pro Phe Lys
335		595	600
337 Pro	Asn Tyr Tyr Arg Tyr		
338		610	

VERIFICATION SUMMARY

DATE: 06/20/2001

PATENT APPLICATION: US/09/782,051

TIME: 14:17:14

Input Set : A:\09-782051.mater.sequencelisting.txt

Output Set: N:\CRF3\06202001\I782051.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:190 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1